



SEQUENCE LISTING

#3

(1) GENERAL INFORMATION:

- (i) APPLICANT: SANCHIS, Vincent
LERECLUS, Didier
MENOUE, Ghislaine
LECADET, Marguerite-Marie
MARTOURET, Daniel
DEDONDER, Raymond
- (ii) TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR
POLYPEPTIDES ENDOWED WITH A LARVICIDAL ACTIVITY TOWARDS
LEPIDOPTERA
- (iii) NUMBER OF SEQUENCES: 2
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
 - (B) STREET: P.O. Box 1404
 - (C) CITY: Alexandria
 - (D) STATE: Virginia
 - (E) COUNTRY: USA
 - (F) ZIP: 22313
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/461,551
 - (B) FILING DATE: 05-JUN-1995
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/251,652
 - (B) FILING DATE: 31-MAY-1994
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 07/458,754
 - (B) FILING DATE: 11-DEC-1989
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: EP 88 401 121.4
 - (B) FILING DATE: 06-MAY-1988
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: FR 87 08090
 - (B) FILING DATE: 10-JUN-1987
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: HUNTINGTON, R. D.
 - (B) REGISTRATION NUMBER: 27,903

(C) REFERENCE/DOCKET NUMBER: 010830-073

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (703) 836-6620

(B) TELEFAX: (703) 836-2021

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2711 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

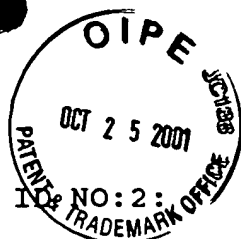
(D) TOPOLOGY: linear

(ii) MOLECULE-TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AAGCTTCAAT AGAATCTCAA ATCTCGATGA CTGCTTAGTC TTTTAAATAC TGTCTACTTG	60
ACAGGGGTAG GAACATAATC GGTCAATTTT AAATATGGGG CATATATTGA TATTTTATAA	120
AATTTGTTAC GTTTTTTGTA TTTTTCATA AGATGTGTCA TATGTATTAA ATCGTGGTAA	180
TGAAAAACAG TATCAAATA TCAGAACTTT GGTAGTTTAA TAAAAAACG GAGGTATTTT	240
ATGGAGGAAA ATAATCAAAA TCAATGCATA CCTTACAATT GTTTAAGTAA TCCTGAAGAA	300
GTACTTTTGG ATGGAGAACG GATATCAACT GGTAATTACT CAATTGATAT TTCTCTGTCA	360
CTTGTTCACT TTCTGGTATC TAACTTTGTA CCAGGGGGAG GATTTTCTAGT TGGATTAATA	420
GATTTTGTAT GGGGAATAGT TGGCCCTTCT CAATGGGATG CATTTCTAGT ACAAATTGAA	480
CAATTAATTA ATGAAAGAAT AGCTGAATTT GCTAGGAATG CTGCTATTGC TAATTTAGAA	540
GGATTAGGAA ACAATTTCAA TATATATGTG GAAGCATTTA AAGAATGGGA AGAAGATCCT	600
AATAATCCAG CAACCAGGAC CAGAGTAATT GATCGCTTTC GTATACTTGA TGGGCTACTT	660
GAAAGGGACA TTCCTTCGTT TCGAATTTCT GGATTTGAAG TACCCCTTTT ATCCGTTTAT	720
GCTCAAGCGG CCAATCTGCA TCTAGCTATA TTAAGAGATT CTGTAATTTT TGGAGAAAGA	780
TTGGGATTGA CAACGATAAA TGTCAATGAA AACTATAATA GACTAATTAG GCATATTGAT	840
GAATATGCTG ATCACTGTGC AAATACGTAT AATCGGGGAT TAAATAATTT ACCGAAATCT	900
ACGTATCAAG ATTGGATAAC ATATAATCGA TTACGGAGAG ACTTAACATT GACTGTATTA	960
GATATCGCCG CTTTCTTTCC AAATATGAC AATAGGAGAT ATCCAATTCA GCCAGTTGGT	1020
CAACTAACAA GGGAGTTTA TACGGACCCA TTAATTAATT TTAATCCACA GTTACAGTCT	1080

GTAGCTCAAT	TACCTACTTT	TAACGTTATG	GAGAGCAGCG	CAATTAGAAA	TCCTCATTTA	1140
TTTGATATAT	TGAATAATCT	TACAATCTTT	ACGGATTGGT	TTAGTGTTGG	ACGCAATTTT	1200
TATTGGGGAG	GACATCGAGT	AATATCTAGC	CTTATAGGAG	GTGGTAACAT	AACATCTCCT	1260
ATATATGGAA	GAGAGGCGAA	CCAGGAGCCT	CCAAGATCCT	TTACTTTTAA	TGGACCGGTA	1320
TTTAGGACTT	TATCAATTCC	TACTTTACGA	TTATTACAGC	AACCTTGCCA	GCGCCACCAT	1380
TTTAATTTAC	GTGGTGGTGA	AGGAGTAGAA	TTTTCTACAC	CTACAAATAG	CTTTACGTAT	1440
GCAGGAAGAG	GTACGGTTGA	TTCTTTAACT	GAATTACCGC	CTGAGGATAA	TAGTGTGCCA	1500
CCTCGCGAAG	GATATAGTCA	TCGTTTATGT	CATGCAACTT	TTGTTCAAAG	ATCTGGAACA	1560
CCTTTTTTAA	CAACTGGTGT	AGTATTTTCT	TGGACGCATC	GTAGTGCAAC	TCTTACAAAT	1620
ACAATTGATC	CAGAGAGAAT	TAATCAAATA	CCTTTAGTGA	AAGGATTTAG	AGTTTGGGGG	1680
GGCACCTCTG	TCATTACAGG	ACCAGGATTT	ACAGGAGGGG	ATATCCTTCG	AAGAAATACC	1740
TTTGGTGATT	TTGTATCTCT	ACAAGTCAAT	ATTAATTCAC	CAATTACCCA	AAGATACCGT	1800
TTAAGATTTC	GTTACGCTTC	CAGTAGGGAT	GCAGCAGTTA	TAGTATTAAC	AGGAGCGGCA	1860
TCCACAGGAG	TGGGAGGCCA	AGTTAGTGTA	GATATGCCTC	TTCAGAAAAC	TATGGAAATA	1920
GGGGAGAACT	TAACATCTAG	AACATTTAGA	TATACCGATT	TTAGTAATCC	TTTTTCATTT	1980
AGAGCTAATC	CAGATATAAT	TGGGATAAGT	GAACAACCTC	TATTTGGTGC	AGGTTCTATT	2040
AGTAGCGTTG	AACTTTATAT	AGATAAAATT	GAAATTATTC	TAGCAGATGC	AACATTTGAA	2100
GCAGAATCTG	ATTTAGAAAG	AGCACAAAAG	GCGGTGAATG	CCCTGTTTAC	TTCTTCCAAT	2160
CAAATCGGGT	TAAAAACCGA	TGTGACGGAT	TATCATATTG	ATCAAGTATC	CAATTTAGTG	2220
GATTGTTTAT	CAGATGAATT	TTGTCTGGAT	GAAAAGCGAG	AATTGTCCGA	GAAAGTCAAA	2280
CATGCGAAGC	GACTCAGTGA	TGAGCGGAAT	TTACTTCAAG	ATCCAAACTT	CAGAGGGATC	2340
AATAGACAAC	CAGACCGTGG	CTGGAGAGGA	AGTACAGATA	TTACCATCCA	AGGAGGAGAT	2400
GACGTATTCA	AAGAGAATTA	CGTCACACTA	CCGGGTACCG	TTGATGAGTG	CTATCCAACG	2460
TATTTATATC	AGAAAATAGA	TGAGTCGAAA	TTAAAAGCTT	ATACCCGTTA	TGAATTAAGA	2520
GGGTATATCG	AAGATAGTCA	AGACTTAGAA	ATCTATTTGA	TCGCGTACAA	TGCAAAACAC	2580
GAAATAGTAA	ATGTGCCAGG	CACGGGTTC	TTATGGCCGC	TTTCAGCCCA	AAGTCCAATC	2640
GGAAAGTGTG	GAGAACCGAA	TCGATGCGCG	CCACACCTTG	AATGGAATCC	TGATCTAGAT	2700
TGTTCCCTGCA	G					2711



(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 823 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Glu	Glu	Asn	Asn	Gln	Asn	Gln	Cys	Ile	Pro	Tyr	Asn	Cys	Leu	Ser	
1			-	5					10					15		
Asn	Pro	Glu	Glu	Val	Leu	Leu	Asp	Gly	Glu	Arg	Ile	Ser	Thr	Gly	Asn	
			20					25					30			
Ser	Ser	Ile	Asp	Ile	Ser	Leu	Ser	Leu	Val	Gln	Phe	Leu	Val	Ser	Asn	
		35					40					45				
Phe	Val	Pro	Gly	Gly	Gly	Phe	Leu	Val	Gly	Leu	Ile	Asp	Phe	Val	Trp	
	50					55					60					
Gly	Ile	Val	Gly	Pro	Ser	Gln	Trp	Asp	Ala	Phe	Leu	Val	Gln	Ile	Glu	
65					70				75						80	
Gln	Leu	Ile	Asn	Glu	Arg	Ile	Ala	Glu	Phe	Ala	Arg	Asn	Ala	Ala	Ile	
			85						90					95		
Ala	Asn	Leu	Glu	Gly	Leu	Gly	Asn	Asn	Phe	Asn	Ile	Tyr	Val	Glu	Ala	
			100				105						110			
Phe	Lys	Glu	Trp	Glu	Glu	Asp	Pro	Asn	Asn	Pro	Ala	Thr	Arg	Thr	Arg	
		115					120					125				
Val	Ile	Asp	Arg	Phe	Arg	Ile	Leu	Asp	Gly	Leu	Leu	Glu	Arg	Asp	Ile	
	130					135					140					
Pro	Ser	Phe	Arg	Ile	Ser	Gly	Phe	Glu	Val	Pro	Leu	Leu	Ser	Val	Tyr	
145					150					155					160	
Ala	Gln	Ala	Ala	Asn	Leu	His	Leu	Ala	Ile	Leu	Arg	Asp	Ser	Val	Ile	
			165					170						175		
Phe	Gly	Glu	Arg	Trp	Gly	Leu	Thr	Thr	Ile	Asn	Val	Asn	Glu	Asn	Tyr	
		180					185						190			
Asn	Arg	Leu	Ile	Arg	His	Ile	Asp	Glu	Tyr	Ala	Asp	His	Cys	Ala	Asn	
		195					200					205				
Thr	Tyr	Asn	Arg	Gly	Leu	Asn	Asn	Leu	Pro	Lys	Ser	Thr	Tyr	Gln	Asp	
	210					215					220					

Trp	Ile	Thr	Tyr	Asn	Arg	Leu	Arg	Arg	Asp	Leu	Thr	Leu	Thr	Val	Leu	225	230	235	240
Asp	Ile	Ala	Ala	Phe	Phe	Pro	Asn	Tyr	Asp	Asn	Arg	Arg	Tyr	Pro	Ile	245	250	255	
Gln	Pro	Val	Gly	Gln	Leu	Thr	Arg	Glu	Val	Tyr	Thr	Asp	Pro	Leu	Ile	260	265	270	
Asn	Phe	Asn	Pro	Gln	Leu	Gln	Ser	Val	Ala	Gln	Leu	Pro	Thr	Phe	Asn	275	280	285	
Val	Met	Glu	Ser	Ser	Ala	Ile	Arg	Asn	Pro	His	Leu	Phe	Asp	Ile	Leu	290	295	300	
Asn	Asn	Leu	Thr	Ile	Phe	Thr	Asp	Trp	Phe	Ser	Val	Gly	Arg	Asn	Phe	305	310	315	320
Tyr	Trp	Gly	Gly	His	Arg	Val	Ile	Ser	Ser	Leu	Ile	Gly	Gly	Gly	Asn	325	330	335	
Ile	Thr	Ser	Pro	Ile	Tyr	Gly	Arg	Glu	Ala	Asn	Gln	Glu	Pro	Pro	Arg	340	345	350	
Ser	Phe	Thr	Phe	Asn	Gly	Pro	Val	Phe	Arg	Thr	Leu	Ser	Ile	Pro	Thr	355	360	365	
Leu	Arg	Leu	Leu	Gln	Gln	Pro	Cys	Gln	Arg	His	His	Phe	Asn	Leu	Arg	370	375	380	
Gly	Gly	Glu	Gly	Val	Glu	Phe	Ser	Thr	Pro	Thr	Asn	Ser	Phe	Thr	Tyr	385	390	395	400
Arg	Gly	Arg	Gly	Thr	Val	Asp	Ser	Leu	Thr	Glu	Leu	Pro	Pro	Glu	Asp	405	410	415	
Asn	Ser	Val	Pro	Pro	Arg	Glu	Gly	Tyr	Ser	His	Arg	Leu	Cys	His	Ala	420	425	430	
Thr	Phe	Val	Gln	Arg	Ser	Gly	Thr	Pro	Phe	Leu	Thr	Thr	Gly	Val	Val	435	440	445	
Phe	Ser	Trp	Thr	His	Arg	Ser	Ala	Thr	Leu	Thr	Asn	Thr	Ile	Asp	Pro	450	455	460	
Glu	Arg	Ile	Asn	Gln	Ile	Pro	Leu	Val	Lys	Gly	Phe	Arg	Val	Trp	Gly	465	470	475	480
Gly	Thr	Ser	Val	Ile	Thr	Gly	Pro	Gly	Phe	Thr	Gly	Gly	Asp	Ile	Leu	485	490	495	
Arg	Arg	Asn	Thr	Phe	Gly	Asp	Phe	Val	Ser	Leu	Gln	Val	Asn	Ile	Asn	500	505	510	
Ser	Pro	Ile	Thr	Gln	Arg	Tyr	Arg	Leu	Arg	Phe	Arg	Tyr	Ala	Ser	Ser	515	520	525	

[illegible]